DOCUMENT RESUME

ED 174 690 TM 009 663

AUTHOR Powers, Stephen
TITLE A Comparison of ESEA Title I Student Gains in a Fall-to-Spring and Fall-to-Fall Out-of-Level Testing Schedule.
PUB DATE Apr 79
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Gains; *Compensatory Education Programs; *Disadvantaged Youth; Educational Trends; Grade 7; Grade 9; High Achievers; Junior High Schools; Longitudinal Studies; Low Achievers; *Reading Achievement; *Student Testing; Testing Programs; *Time Factors (Learning)
IDENTIFIERS Comprehensive Tests of Basic Skills; Elementary Secondary Education Act Title I; Metropolitan Achievement Tests; *Out of Level Testing; *Test Score Decline

ABSTRACT

A longitudinal study was initiated in the Tucson Unified School District (TUSD) to follow the achievement of seventh and ninth grade Title I students: (1) to investigate the suitability and difficulty of the Total Reading Test of the Metropolitan Achievement Test (MAT) and the Comprehensive Tests of Basic Skills (CTBS); (2) to compare the gains in a fall-to-spring out-of-level testing program; (3) to determine if there was a decline in reading achievement over the summer; and (4) to study the fall-to-fall reading achievement patterns of low- and high-achieving groups in a compensatory education program. Results of the study indicated: the CTBS was suitable for both groups; the MAT was suitable for the ninth grade group but unsuitable for the seventh grade group; the seventh grade group made greater gains in fall-to-fall testing and the ninth graders made greater gains in fall-to-spring testing; the ninth graders but not the seventh graders showed a decline in reading achievement over the summer; and homogeneous regression lines of low- and high-achieving groups were retained. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED174690

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Stephen Powers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A Comparison of ESEA Title I Student Gains
in a Fall-to-Spring and Fall-to-Fall
Out-of-Level Testing Schedule

Introduction

David and Pelavin (1978) have argued that fall-to-spring evaluations may be misleading because over the summer months, students suffer large losses in achievement. When Title I evaluators administer norm-referenced achievement tests one level below that recommended by the test publisher, further suspicions arise as to whether the level of achievement attained in the spring will be maintained across the summer months.

The loss in achievement often found over the summer months may be due to the use of inappropriate tests with compensatory students. Using tests one level below that recommended by the test publisher is supposed to be a more accurate measure with compensatory education students. Roberts (1976) pointed out that administering an out-of-level test is not always appropriate for compensatory education students. However, when out-of-level testing is warranted, it may assist in avoiding floor effects which occur when a test which is inordinately difficult, is administered to low-achieving students. Moreover, out-of-level testing may provide a more reliable measure of the ability of low-achieving students (Roberts 1976).

Researchers have disagreed about the need for out-of-level testing with compensatory education students. Wick (1973, 1977), Powers (1977) and Haenn and Proctor (1978) recommended out-of-level testing.

Slaughter and Gallas (1978) considered out-of-level testing as merely a stopgap measure. McNamara (1973), Long, Schaffran, Ayrer and Kellog (1977) and Ozenne (1978) have identified problems in cross-form and vertical scaling.

In the 1977-78 school year, the Elementary and Secondary Education Act (ESEA) Title I compensatory reading project was provided to approximately 4300 students in grades K-12 of the Tucson Unified School District (TUSD). In the fall of 1977, a longitudinal study was initiated to follow the achievement of two cohorts of Title I students across several years of Title I instruction. The purposes of the present study are: 1) to investigate the suitability and difficulty of the Total Reading Test of the Metropolitan Achievement Test (MAT) and the Comprehensive Tests of Basic Skills (CTBS), 2) to compare the gains in a fall-to-spring and fall-to-fall out-of-level testing schedule, 3) to determine if there was a decline in reading achievement over the summer months, and 4) to study the trend in achievement from fall to fall to discover if there were different achievement patterns between a low and a high achieving group in a compensatory education program.

Methodology

The present study employed a longitudinal single--group design with a seventh grade cohort and a ninth grade cohort. All Title I students in the TUSD junior high schools and high schools were tested in the fall 1977, spring 1978 and fall 1978 with the CTBS, Form S, Reading Vocabulary and Reading Comprehension Tests. Seventh grade students were administered level 2 of the CTBS, and ninth grade students were administered level 3. These levels were one level below the level which the test publishers recommended. Title I students were also tested

in the fall 1977 with the MAT, Form F, Advanced, Word Knowledge and Reading subtests. The research design is presented in Figure 1.

- - - - - - - - - - - - - - - -
Insert Figure 1 about here
- - - - - - - - - - - - - - - -

The CTBS, Levels 1-4, is a battery of seven tests measuring three basic skills areas: Reading, Language and Mathematics (CTBS Test Coordinator's Handbook 1976). The skills areas were classified using Bloom's Taxonomy of Educational Objectives. In the test development, efforts were made to reduce racial and ethnic bias. The K-R 20 reliabilities at each grade level for vocabulary, comprehension and total scores are almost all above .90 with standard errors of measurement from .25 and 1.01 in grade equivalent (GE) units. Moreover, it appeared that systematic procedures were followed in test development to ensure content validity. The CTBS Reading Comprehension subtests, Levels 1-4 are composed of 45 items and each item in Levels 1-4 contains a multiple choice involving four alternatives. The Reading Comprehension: Passages subtest in Level C included 18 items, each item with a multiple choice of four alternatives. For an outline of the grade levels recommended for administration of the CTBS, refer to Table 1.

- - - - - - - - - - - - - - - -
Insert Table 1 about here
- - - - - - - - - - - - - - - -

Two cohorts were formed of all seventh and ninth grade students who were enrolled in the Title I project in the TUSD and who were tested with the CTBS in the fall of 1977. Attrition was the major design problem in this study. Of the 698 seventh and ninth grade students forming the two cohorts in the fall of 1977, 231 were present for testing

in the spring of 1978 and the fall of 1978. This represented 33 percent of the original group. See Table 2.

- - - - - - - - - - - - - - - - -
Insert Table 2 about here
- - - - - - - - - - - - - - - - -

With such a large reduction in sample size, the generalizability of this study was severely limited. Therefore, no attempt is made to generalize the findings to the 698 students in the initial group of students; rather, the results of this study will be interpreted in light of those students who participated in the Title I project from fall 1977 to fall 1978.

The suitability and difficulty of the MAT and CTBS was investigated using guidelines advocated by Roberts (1976). Fall-to-spring and fall-to-fall evaluations were accomplished with the Norm-Referenced Model A1 (Tallmadge and Wood, 1976). Dependent t-tests were used to compare the spring 1978 mean with the fall 1978 mean of each cohort. To study the achievement trends of low and high achieving groups in each cohort, simple regression analyses were used. Interaction between the regression lines of the low and high groups was studied using a procedure described by Kerlinger (1973).

Before data analysis, CTBS raw scores were converted to the CTBS Expanded Standard Score Scale. Calculations were performed with expanded standard scores unless otherwise noted. As the CTBS Expanded Standard Score Scale is a normalized scale with assumed equal intervals, it was believed this scale was more appropriate for statistical analysis. Moreover, the CTBS Expanded Standard Score Scale is purported to measure the underlying reading comprehension achievement dimension (CTBS Technical Bulletin No. 1, (1974) page 10). The choice of this scale is

The mean reading ability in spring 1978 was compared with the mean reading ability of the fall 1978. A significant difference was found for the seventh grade cohort ($p < .001$); this indicated the ability of the seventh grade cohort in the fall 1978 to be superior to its level of attainment of the spring 1978.

The ninth grade cohort's mean reading ability of the spring 1978 was compared with the fall 1978 mean reading ability. This t-test yielded nonsignificant results. Therefore, there was no evidence to consider the level of achievement in the fall significantly different from that of the spring. Refer to Table 6.

- - - - - - - - - - - - - - - -
Insert Table 6 about here
- - - - - - - - - - - - - - - -

When mean expanded scale scores were converted to NCEs, it appeared the ninth grade students declined from 24.5 NCEs in the spring to 22.8 NCEs in the fall. The gains of the seventh grade cohort appeared small when expressed in NCEs. The seventh grade cohort gained from 26.1 NCEs in the spring to 27.9 NCEs in the fall. Since NCEs are linked with percentile ranks of the norm group, it appeared the ninth grade cohort declined over the summer with respect to the national norms.

The trends in reading achievement of low-achieving and high-achieving Title I students were studied. Each cohort was divided into a high and low group at the median based on the fall 1977 Total Reading Score of the CTBS. The fall 1978 CTBS scores were regressed on the fall 1977 scores in order to obtain an estimate of the students' trend in achievement. Then the regression lines were compared using multiple-regression analysis (Kerlinger 1974) to determine if there was any interaction. No significant interaction was found. Therefore, the

hypothesis of homogeneity of regression was retained for the comparisons of the high and low groups in the seventh grade cohort and the high and low groups of the ninth grade cohort. Refer to Table 7.

- - - - - - - - - - - - - - - -
Insert Table 7 about here
- - - - - - - - - - - - - - - -

Summary and Discussion

The purposes of the present study were: 1) to investigate the suitability and difficulty of the Total Reading Test of the MAT and CTBS, 2) to compare the gains in a fall-to-spring and a fall-to-fall out-of-level testing schedule, 3) to determine if there was a decline in reading achievement over the summer months, and 4) to study the trend in achievement from fall to fall to discover if there were different achievement patterns between a low and high achieving group.

The suitability and difficulty of the MAT and CTBS Total Reading Test were investigated in this study as a preliminary to further data analysis. According to the guidelines for test suitability (Roberts 1976), the CTBS appeared to be suitable for the seventh and ninth grade cohorts. The MAT did not appear to be suitable for the seventh grade cohort; however, the MAT seemed suitable for the ninth grade cohort. No floor effects were apparent for the seventh or ninth grade cohorts on either the MAT or CTBS.

Neither the seventh grade cohort nor the ninth grade cohort showed decline from spring 1978 to fall 1978 when measured in expanded standard scores. The evaluation of gains from fall-to-spring and fall-to-fall was accomplished with NCEs (Tallmadge and Wood 1976). From fall 1977 to spring 1978 the seventh grade cohort gained 4.9 NCEs and the ninth grade cohort gained 2.6 NCEs. From fall 1977 to fall 1978

the seventh grade cohort gained 6.7 NCEs and the ninth grade cohort gained .9 NCEs. In summary, the seventh grade cohort made greater gains fall-to-fall than fall-to-spring, whereas the reverse was true for the ninth grade cohort.

The major question of this study was whether compensatory education students would decline in achievement over the summer months when they are tested out-of-level. Two scales were used to compare the achievement of the seventh and ninth grade cohorts across the summer: The CTBS Expanded Standard Score Scale which purports to measure the underlying reading achievement dimension and the NCE scale which measures achievement relative to the norm group's percentile rank.

When measured on the CTBS Expanded Standard Score Scale, the seventh grade cohort appeared not only to maintain its spring 1978 achievement level, but to achieve significantly higher in the fall 1978 ($p < .001$). The ninth grade cohort seemed to maintain its spring 1978 achievement level across the summer to fall 1978.

When mean expanded scale scores were converted to NCEs, it appeared the seventh grade cohort gained 1.8 NCEs from spring 1978 to fall 1978, and the ninth grade cohort declined--1.7 NCEs between spring and fall. The apparent decline of the ninth grade cohort over the summer months was such that the fall 1978 level was almost equal to the fall 1977 level.

There was no evidence to support the hypothesis that the seventh grade or ninth grade cohorts achieved significantly lower than in the spring 1978 level as measured with expanded standard scores. The seventh grade cohort actually gained significantly over the summer months. NCEs provided another interpretation for student gains because NCEs are

linked to the percentile ranks of the norm group (Tallmadge and Wood 1976). The ninth grade cohort appeared to decline in achievement over the summer. The decline was almost to the level of fall 1977.

In each cohort, the trends in reading achievement of a low group and a high group were studied. Regression lines were compared for the purpose of discovering any significant interaction. No significant interaction was discovered; therefore, the hypothesis of homogeneous regression lines was retained.

Conclusions

This study extended the research in summer loss to the developing area of out-of-level testing. Summer loss was not apparent with the seventh grade cohort but it did appear with the ninth grade cohort when achievement was measured on the NCE scale.

The results of the analysis of the seventh grade and ninth grade cohorts using the CTBS Expanded Standard Score Scale did not support the hypothesis of summer loss in achievement. Neither the seventh nor ninth grade cohorts gave evidence of decline over the summer. On the contrary, the seventh grade cohort demonstrated significant gain over the summer.

When the seventh and ninth grade cohorts' achievement was measured with the NCE scale, the ninth grade cohort appeared to decline over the summer. The decline was such that the achievement level of fall 1978 was almost the level of fall 1977. The seventh grade cohort did not show a decline over the summer months, even when measured on the NCE scale. The ninth grade cohorts' results measured on the NCE scale supported findings of David and Pelavin (1978) that there is decline in achievement over the summer months.

REFERENCES

Ayrer, J. E., & McNamara, T. C. Survey testing on an out-of-level basis. Journal of Educational Measurement, 1973, 10(2) 79-84.

CTBS Technical Bulletin No. 1, Form S, Expanded Edition, 1974, Monterey, California.

David, J. D. Pelavin, S. H. Evaluating Compensatory Education: Over What Period of Time Should Achievement be Measured? Journal of Educational Measurement, 1978, 15(2), 91-99.

Haenn, J. F., & Proctor, D. C. A practitioner's guide to out-of-level testing. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada, March 27-31, 1978.

Long, J. V., Schaffran, J. A., & Kellogg, T. M. Effects of out-of-level survey testing on reading achievement scores of Title I, ESEA students. Journal of Educational Measurement, 1977, 14(3), 203-213.

Ozenne, D. The effect of vertical scaling imprecision in the estimation of Title I project gains. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada, March 27-31, 1978.

Powers, S. Implications of out-of-level testing for ESEA Title I students. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada, March 27-31, 1978.

Roberts, O. Out-of-level testing: ESEA Title I evaluation and reporting system. RMC Research Corporation, Mountain View, October 1976.

Slaughter, H. B. & Gallas, E. J. Will out-of-level norm-referenced testing improve the selection of program participants and the diagnosis of reading comprehension in ESEA Title I programs? Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada, March 27-31, 1978.

Tallmadge, G. K. and Wood, C. T. User's Guide: ESEA Title I Evaluation and Reporting System. RMC Research Corporation, Mountain View, California, December 1976.

Wick, J. W. Educational Measurement, Merrill Publishing Company, Columbus, Ohio, 1973.

Wick, J. W. A report on individualized testing in Chicago. Putting Current Ideas in Measurement to Work in the Schools. Proceedings from the Thirteenth Annual Houghton-Mifflin Measurement Conference, Iowa City, Iowa, April 26-27, 1977.

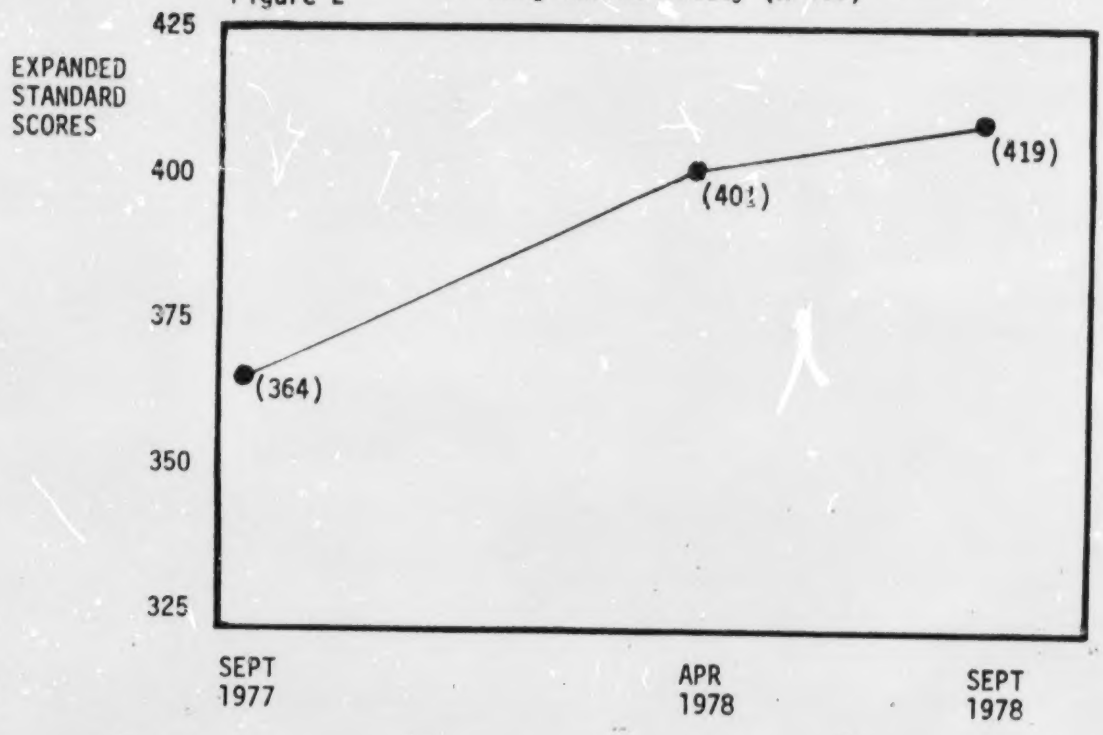
Figure 1. Longitudinal Single-Group Design of Title I Students Tested with the Comprehensive Tests of Basic Skills (CTBS), Total Reading Test and Metropolitan Achievement Test (MAT), Total Reading Test

Cohort	Fall 1977	Spring 1978	Fall 1978
Seventh Grade N = 150	CTBS MAT	CTBS	CTBS
Ninth Grade N = 81	CTBS MAT	CTBS	CTBS

ESEA TITLE I SEVENTH GRADE COHORT

Longitudinal Study (N=159)

Figure 2



ESEA TITLE I NINTH GRADE COHORT

Longitudinal Study (N=81)

Figure 3

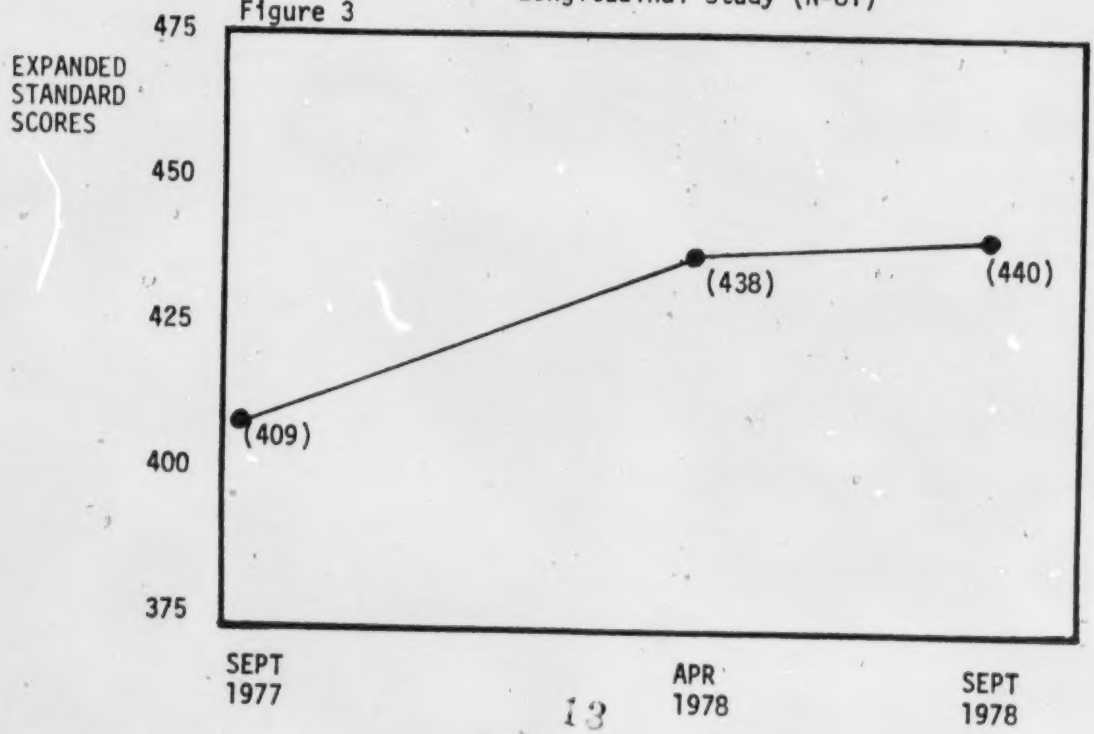


Table 1. Comprehensive Tests of Basic Skills (CTBS) Test Levels and Recommended Grades

Test Level	Grades
A	K.0 - 1.3
B	K.6 - 1.9
C	1.6 - 2.9
1	2.5 - 4.9
2	4.5 - 6.9
3	6.5 - 8.9
4	8.5 -12.9

Table 2. Percent of Title I Students with Comprehensive Tests of Basic Skills (CTBS) Test Scores in Fall 1977, Spring 1978, and Fall 1978

Cohort	Fall 1977	Spring 1978		Fall 1978	
	N	N	Percent	N	Percent
Seventh Grade	370	314	85	150	41
Ninth Grade	328	219	67	81	25
TOTAL	698	533	76	231	33

Table 3. The Suitability of the Total Reading Test of the CTBS and MAT.*

Grade	Test	N**	$\bar{x}$	Maximum Score	Interval***	Suitable?
Seventh	MAT	344	23.6	95	31.7 - 71.2	No
	CTBS	370	32.6	85	28.3 - 63.8	Yes
Ninth	MAT	240	33.4	95	31.7 - 71.2	Yes
	CTBS	328	31.0	85	28.3 - 63.8	Yes

* Statistics have been computed from raw scores.

** Number of students tested in the fall of 1977.

*** The level of a test is suitable when the raw score is equal to or above one-third of the maximum score and somewhat less than three-quarters of the maximum score (Roberts 1976).

Table 4. Floor Effects in the Total Reading Test of the CTBS and MAT.*

Grade	Test	N**	$\bar{x}$	Md	$\bar{x}$-Md***	1/3 SD
Seventh	MAT	344	23.6	23.6	.0	2.4
	CTBS	370	32.6	31.0	1.6	3.6
Ninth	MAT	240	33.4	32.4	1.0	3.8
	CTBS	328	31.0	30.1	.9	4.27

* The statistics have been computed from raw scores.

** Number of students tested in the fall of 1977.

*** If the mean is higher than the median by about one-third of a standard deviation, a floor effect may have been encountered (Roberts 1976).

Table 5. Mean Normal Curve Equivalents (NCEs), CTBS Expanded Standard Scores (ESS) and Grade Equivalents (GEs) for Seventh and Ninth Grade Cohorts.

Cohort	Fall 1977			Spring 1978			Fall 1978		
	NCE	ESS	GE	NCE	ESS	GE	NCE	ESS	GE
Seventh Grade N=150	21.2	364	3.4	26.1	401	4.2	27.9	419	4.6
Ninth Grade N=81	21.9	409	4.4	24.5	438	5.2	22.8	440	5.2

Table 6. Comparison of the Spring 1978 and Fall 1978 Achievement Levels of the Seventh and Ninth Grade Cohorts

Cohort	N	$\bar{x}_1$	Standard Error	$\bar{x}_2$	Standard Error	Correlation	t
Seventh Grade	150	401.4	4.7	419.4	4.3	.76	5.65*
Ninth Grade	81	437.7	7.8	440.0	9.1	.49	.26

* $p < .001$

Table 7. Trends in Reading Achievement of Low Achieving and High Achieving Students

Cohort	Level	N	B	Constant
Seventh Grade	Low	71	.76	140.0
	High	79	.80	130.5
Ninth Grade	Low	42	.61	178.5
	High	39	.47	266.2

APPENDIX A

Appendix A

Descriptive Statistics of Seventh and Ninth Grade Students Tested in Fall 1977, Spring 1978 and Fall 1978 with the Comprehensive Tests of Basic Skills, Total Reading Test

Cohort	N	Fall 1977 $\bar{x}$	Fall 1977 SD	Spring 1978 $\bar{x}$	Spring 1978 SD	Fall 1978 $\bar{x}$	Fall 1978 SD
Seventh Grade	370*	376.5	51.5				
	314**	373.0	49.3	410.6	60.2		
	150***	363.8	44.8	401.4	58.2	419.4	52.9
Ninth Grade	328*	419.2	77.5				
	219**	421.4	68.6	441.5	85.0		
	81***	409.4	56.2	437.7	70.2	440.0	82.1

* All Title I students tested in the fall 1977

** All Title I students with scores in fall 1977 and spring 1978

*** All Title I students with scores in fall 1977, spring 1978 and fall 1978